DOCUMENT RESUME

ED 436 548 TM 028 378

AUTHOR Impara, James C.; Plake, Barbara S.; Hertzog, Melody; Giraud, Gerald; Spies, Robert

TITLE Utility of a Concept-Focusing Strategy on Judgmental Standard Setting Results.

PUB DATE 1998-04-00

NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998). For closely related paper, see ED 417 231.

PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Concept Formation; *Decision Making; *Evaluation Methods; *Judges; Performance Factors; *Standards; Validity

IDENTIFIERS Experts; *Standard Setting

ABSTRACT

Judgmental standard setting approaches rely on the perceptions of experts about examinee performance on a test. Traditional standard setting methods ask panelists to estimate how well a randomly selected hypothetical examinee who is representative of a well-defined target group, usually a minimally competent candidate (MCC), will perform on each item on a test. Item performance estimates are difficult for panelists to make accurately; however, the validity of these performance standards rests on the ability of the panelist to conceptualize the skills and abilities of the MCC accurately and make accurate performance estimates. This study investigated the utility of a strategy to aid in the conceptualization of the MCC. Panelists were asked to envision the MCC and the typical student, and then make performance estimates, first for the typical student and then for the MCC. Results showed that the strategy resulted in significantly lower performance standards than did the traditional approach. Validity data were used to evaluate the accuracy of the judgments resulting from the experimental and traditional approaches. More research is needed to clarify the utility of concept-focusing strategy on the judgmental performance standards. (Contains 2 tables and 10 references.) (Author/SLD)

# Utility of a Concept-Focusing Strategy on Judgmental Standard Setting Results

James C. Impara, Barbara S. Plake, Melody Hertzog,
Gerald Giraud, Robert Spies
University of Nebraska-Lincoln



Running head: Concept-focusing strategy



Paper presented at the 1998 Annual Meeting of the American Educational Research Association, San Diego, CA.

# Utility of a Concept-Focusing Strategy on Judgmental Standard Setting Results

## Abstract

Judgmental standard setting approaches rely on the perceptions of experts about examinee performance on a test. Traditional standard setting methods ask panelists to estimate how well a randomly selected hypothetical examinee who is representative of a well defined target group, usually a minimally competent candidate (MCC), will perform on each item on a test. Item performance estimates are difficult for panelists to make accurately; however the validity of these performance standards rests on the ability of the panelists to accurately conceptualize the skills and abilities of the MCC and make accurate performance estimates. This study investigated the utility of a strategy to aid in the conceptualization of the MCC. Results showed that the strategy resulted in significantly lower performance standards than did the traditional approach. Validity data were used to evaluate the accuracy of the judgments resulting from the experimental and traditional approaches. More research is needed to clarify the utility of a concept-focusing strategy on the judgmental performance standards.

# Utility of a Concept-Focusing Strategy on Judgmental Standard Setting Results

## Introduction

Typically, the purpose of standard setting procedures is to separate examinees into multiple groups by setting a performance standard, or cut score, that can be used to differentiate between group members. Usually there are two groups: masters and nonmasters, but the number of groups may vary. This cut score is then used as the criterion to determine which examinees are assigned to which groups. In many settings, the consequences of being classified as a nonmaster can be severe: licenses or certifications may be denied or students may be denied graduation or promotion in school settings. Therefore, the accuracy of these performance standards can be critically important to individual examinees in terms of their future admission in professions or other educational programs.

Setting cut scores is a policy decision (Kane, 1994). Such decisions are often informed by the results of standard setting studies that involve using panelists to make judgments about how examinees will perform on test items (judgmental standard setting methods), or that involve panelists making classification decisions about examinees who are known to them (examinee-centered standard setting methods), or using some empirical strategy (usually involving a regression model).

In judgmental standard setting methods, panelists are typically asked to estimate, or predict, the performance on the test questions by examinees who are on the borderline between two classifications. For simplicity, we discuss the two-group case only. In the two-group case this borderline group of examinees is referred to as the Minimally Competent Candidates (MCCs). These examinees represent the group of candidates whose knowledge, skills, and achievements are just barely sufficient to warrant a "passing" score. In the judgmental standard setting method proposed by Angoff (1971) panelists estimate the performance on each item in the test by a randomly selected, hypothetical MCC. Often this task is operationalized as estimating how many out of 100 hypothetical MCCs will correctly answer each test question. In other applications of the Angoff method, panelists make a dichotomous choice about whether a known MCC will answer each item correctly or not (Impara & Plake, 1997).

The item performance estimates form the basis for determining the recommended performance standard. Most often, each panelist's item performance estimates are summed across the items in the test to set the individual panelist's cut score. These cut scores are then averaged across panelists to determine the recommended cut score, or performance standard, for the test.

Therefore, when panelists estimate the probability of a hypothetical MCC answering correctly, the resulting performance standards from the judgmental standard setting methods rely directly on the accuracy of the item performance estimates made by the panelists. These item performance estimates are actually conditional probability estimates (i.e., they represent the panelists' estimate of correct performance conditioned on the ability level of the MCC). In essence, the panelists are estimating item difficulties for a specific target group of examinees, the MCCs. The panelists' perception of the skills and competencies of the MCC, then, are central to the accuracy of these item performance estimates.

Research has shown that item performance estimates are difficult for panelists to make accurately. Even when the target group of candidates is the total group, panelists have difficulty providing accurate performance estimates

(Lorge & Kruglov, 1953, Thorndike, 1982; Bejar, 1983). In a study focusing on the just competent student in sixth grade science, Impara and Plake (1998) found that teachers, who were familiar not only with their students but also with the assessment, systematically underestimated the performance of their "D/F" students (which served as the operational definition of the minimally competent student in their classes). Therefore, a strategy that improved the accuracy of item performance estimates would make an important contribution to the methodology of judgmental standard setting approaches.

The purpose of this study was to investigate the utility of a concept-focusing strategy on the performance standards set using a judgmental standard setting approach. Formation of a clear concept of the MCC is critical to the accuracy of the judgments. The concept is generally developed through group discussion of the characteristics of such examinees, focused on the MCC's performance on the objectives measured by the assessment. It may be possible to improve the clarity of the concept in the minds of the panelists by contrasting it with a familiar, related concept. In this case, the panelists were instructed to envision two examinee groups: the MCC and the typical student. Panelists made item performance estimates for the typical student first, then made performance estimates for the MCC. The impact of this concept-focusing strategy was investigated in this study. This method is being used in at least one certification testing program (Goodwin, in press), but there has been no evidence reported to support the validity of the method.

<u>Method</u>

A large midwestern school district served as the setting for this study. This school system has adopted a Gateway Assessment Program aimed at identifying students from differing developmental levels in a variety of content areas who are in need of additional educational services. For each content area, Essential Learner Outcomes (ELOs) have been established. The district uses a series of assessments to measure student achievement tied to these ELOs; minimum passing scores are determined, using a judgmental standard setting approach, to identify students who qualify for additional services. The High School Mathematics Gateway Assessment was used for this investigation.

Instrument. The High School Mathematics Gateway Assessment consists of 62 questions, 30 items selected from the Metropolitan Achievement Test (MAT) and 32 items (comprising the Supplemental Test) developed specifically for this assessment to measure the ELOs for high school mathematics not measured by the MAT. Items from the MAT were exclusively multiple choice, whereas items on the Supplemental Test were primarily short-answer, constructed response items. The Supplemental Test items are scored using a pre-established rubric by trained assessors; the key provided by the MAT was used to score the MAT items. The assessment was administered to all students in 10th grade during the previous semester. No students in grades other than grade 10 participated in the assessment program. Student scores had not been reported at the time of the standard setting workshop.

Panelists. A total of 24 panelists participated in the study, which was administered as part of the operational standard setting workshop for high school mathematics. All panelists were high school mathematics teachers in the school system. The teachers represented a variety of content areas within mathematics and taught students at varying levels of mathematics, from remedial mathematics to pre-calculus. All of the high schools in the district were represented on the panel. The panelists were divided into two groups in such a way that there was a balance in the schools represented and other relevant factors such as teaching experience and specific mathematics class taught.

Procedure. Prior to participating in the standard setting workshop, most of the panelists were asked to make "global ratings" of their students on their projected performance on the High School Mathematics Gateway Assessment. Using only the table of specification for the assessment as the frame of reference, these teachers were asked to classify each of their 10th grade students into one of 4 performance categories: NO: have not mastered sufficient skills to be deemed masters; SOME: have mastered some of the skills to be deemed masters; YES: have mastered sufficient skills to be considered masters, and BORDERLINE: have just the minimum level of skills to be at the SOME level.

All panelists were in a single group at the outset for the orientation and training. The orientation consisted of a description of purpose of the standard setting workshop. The panelists were all familiar with the High School Mathematics Gateway Assessment, but the table of specifications was reviewed nonetheless. After a discussion of the assessment, the panelists engaged in a concept-formation exercise where they were first asked to think of a specific student they felt was "just barely competent" in the ELOs measured by the test. By describing the knowledge, skills, and achievements of those students, the group arrived at a definition of the Just Competent Student (JCS) analogous to the MCC. This discussion was directed specifically at the content components of the High School Mathematics Gateway Assessment. At the conclusion of this discussion, the panelists were divided into two groups, meeting in separate rooms.

After separating into two groups, training continued. The training for both groups consisted of making item performance estimates on a set of test items similar to those used in the assessment. These items were modeled after the MAT items and had been administered to a sample of 10th grade students in the school system. For each item in the short practice test, panelists were asked to make independent judgments of whether the JCS they had in mind during the earlier discussion would answer this question correctly. This is a slight modification of the traditional Angoff Standard Setting Method where panelists are asked to make item performance estimates using the full probability scale. This variation, called the Yes/No Method (Impara & Plake, 1997) has been shown to provide comparable results to the Angoff approach in less time. After Round 1 in the practice session, panelists discussed their reasons for believing the JCS would answer each practice question in a particular way. This discussion tended to further clarify the definition of the JCS and to help panelists connect the more abstract characteristics of a JCS with performance on a specific test item. There was no discussion of this type during the operational standard setting.

In Group 1, the control group, after completing their initial item performance decisions, panelists were shown data on how students in the school system performed on the test during the most recent administration. Data included the proportion of the total group of students who answered each item correctly (p-values for each item) and the impact of employing the panelists' Round 1 cutscore on the proportion of students system-wide who would qualify for additional educational services. Panelists were given item performance information for each of the four performance groups (NO, BORDERLINE, SOME, YES) of students as well. Following discussion of these data, panelists were given the opportunity to revise their initial item performance decisions; these second ratings were their round 2 estimates.

Group 2, the experimental group, followed the steps outlined above for Group 1 with one exception. Panelists in the experimental group were asked to make item performance decisions for two levels of student performance: the "typical" student followed by the JCS. Panelists were told to consider the typical student for the district as a whole, which may differ from the typical student in the courses they currently teach. As with the control group, panelists had an opportunity to practice making their judgments and interpreting empirical data. However, their discussion during the practice session involved reasons for the judgments they had made about the typical student as well as those made for the JCS.

Results

The objective of this study was to investigate if the use of a concept focusing strategy would assist panelists in making more accurate performance estimates for examinees classified as Just Competent. Specifically, if panelists were asked to estimate performance for the "typical" student and then provided feedback, would this help the panelists to focus on the target students of interest, the JCS group.

The planned analyses consisted of examining the accuracy of the panelists in the experimental group in terms of estimating the performance of the typical student, as represented by the difficulty level (p-value) of the total group of students for both rounds one and two. Moreover, the extent that the performance estimates changed, both the means and variances, between rounds one and two for the typical students and for the JCS, would provide insights about the influence of the concept focusing strategy. If the panelists in the two groups differed significantly in their round two estimates of the JCS, then a comparison of the item performance estimates from both groups of panelists would be compared with the actual performance of students classified as BORDERLINE by the teachers.

The best laid plans sometimes go awry. When the time came for conducting these analyses, it was impossible to do what had been planned. The data had been lost when the server crashed. The only data available were selected data that had been printed out or saved on another hard disk. Consequently, the analyses reported below reflect the initial intentions, but with some components missing.

The first analysis consisted of examining the degree to which panelists in the experimental group were able to estimate known values: the actual performance of the typical student. In Table 1, the round 2 performance estimates of typical students for the experimental group panelists was 54.34. The actual performance of the typical student was 42.36. Thus, the panelists overestimated the performance of the typical student by about 12 score points on the 62 item test. If the objective was to set a cut score for this type of student (when the actual value was known) the panelists would have done a very poor job of estimating actual difficulty. The panelists were able, however, to rank order the items with some degree of accuracy as the correlation between the panelists' item performance estimates and the actual p-values for the typical student (the total group) was .793.

The recommended cut scores, from each group of panelists, for the JCS, based on each group's second round Angoff performance estimates, are shown in Table 1. The cut score derived from the control group's estimates equaled 36.08 (SD = 6.40), compared to 28.25 (SD = 6.46) for the experimental group. These values differ significantly, ($t(22) = 2.9$; $p < .01$).

****Table 1 about here****

To try to gain insights about the validity of the divergent performance estimates from these two panels, their estimates were compared to the p-values for the students classified as BORDERLINE by the teachers who made global ratings of their students' competency in mathematics. The RMSE of the estimates was computed for each group across all 62 items. Neither group was particularly consistent with the actual performance of students classified as BORDERLINE.

As shown in Table 1, the control group's RMSE equaled 0.236 while the RMSE for the experimental group's RMSE value was 0.252. These values, though, differ systematically in the direction of the error: the performance estimates of the panelists in the control group was systematically higher than the actual performance of students classified as BORDERLINE (an absolute difference of 1.91 points), whereas panelists in the experimental group made estimates that were systematically lower than the actual performance of the BORDERLINE group (-5.94 points).

To examine if the panelists' estimates were at least consistent with the actual difficulty levels of the items, correlations were computed between the panelists' performance estimates and the actual item difficulties. These correlations, shown in Table 2, suggest that the panelists were able to judge relative difficulty with some degree of accuracy. The lowest correlation was for panelists in the control group whose correlation between performance estimates and the actual item difficulties for students classified as BORDERLINE was: $r = .608$. The highest correlation was between the round two performance estimates for borderline students and actual performance for BORDERLINE students from panelists in the experimental group: $r = .867$.

****Table 2 about here****

Conclusions and Discussion

Panelists in the control and experimental groups produced performance standards that differed significantly. One element of the validity data, the comparison between the performance estimates for the typical students from the panelists in the experimental group was extremely inaccurate. If these data were used in the same was as Goodwin (in press) used them -- as an adjustment to the estimate of the MCC -- the correction would be substantial and of questionable validity. Other validity data comparing the performance estimates with actual data for the BORDERLINE students, suggest that neither group's performance estimate was superior in accuracy; the groups differed from the target group's performance by different amounts and in opposite directions. However, the global ratings are themselves fallible as validity measures.

Two different explanations might provide different bases for questioning the accuracy of the mean score of the borderline group. First, it is possible that global ratings are influenced by a halo effect to a greater extent than are item-level ratings. This might take the form of a systematic negative bias in global ratings of students who have not mastered several of the test's objectives, a tendency to generalize low expectations of their performance. If this were occurring, then some students who belong in the SOME category would probably be erroneously classified as BORDERLINE. This could inflate the p-values for the BORDERLINE group, which would change the conclusion drawn concerning the accuracy of the panelists' estimates. In such a case, the estimates of the experimental group would be more accurate than those produced by the control group.

To examine this possibility, other data were examined along with the actual scores of the students classified as borderline. The other data included the mathematics course in which the student was enrolled. Some of the students classified as BORDERLINE were enrolled as 10th graders in a precalculus class. Because the test measured content covered up to the first semester in beginning algebra, we did not expect any precalculus students to be classified as BORDERLINE. We believe that some teachers classified students based on the student's performance in the teacher's class, not on the student's ability to correctly answer the items on the test. The one precalculus student classified as BORDERLINE answered 57 items correctly (out of 62). This suggests that there is some credence to the hypothesis that some teachers misunderstood the basis for the classification of students.

The modal student classified as BORDERLINE was enrolled in geometry. Most of these geometry students obtained scores of 40 or more. There were, of course, some much lower scores by students enrolled in geometry, and by students enrolled in algebra classes. Virtually none of the students who were enrolled in prealgebra classes were identified as being borderline (most were classified in the NO category). Virtually all of the students in the latter classes obtained scores of less than 30 correct. The foregoing hypothesis is purely speculative, but illustrates the need for further investigation of the accuracy of the global ratings and identification of additional sources of validity information.

The second explanation that may provide insights about the adequacy of the BORDERLINE group mean score as the basis for comparing the Angoff ratings is found in Livingston (1995). Livingston suggests that the borderline group method is subject to regression to the mean. Because the group is identified as a low scoring group based on a measure (teacher ratings) that is correlated with the criterion measure (the test), the students classified at the extremes are subject to the regression effect. This, too, would result in an overestimation of the cut score for the students classified as BORDERLINE, making the experimental group's cut score estimate the more acceptable.

One limitation of the current study was the absence of a clear definition of the typical student. After Round 1, the variance of individual panelists' estimates of the minimum passing score was approximately four times as large for the experimental group as for the control group. As expected, the variances across the two groups of judges converged during Round 2, after feedback about actual student performance was given. One plausible explanation for this finding is that initially the experimental group was not uniform in its interpretation of "typical," but that the empirical data helped to create a common definition. In future studies, a definition should be developed explicitly as part of the panelists' training in the same way that a definition of the JCS is developed.

Strategies to improve judgmental standard setting would benefit from a deeper understanding of the mental process a panelist goes through as he or she conceptualizes the JCS and makes a performance estimate. To some extent, "just competent" only has meaning if we have a concept of "competent." To what extent do panelists implicitly contrast the concept of the JCS to a typical student? Alternatively, do they apply some other standard of comparison as they define the JCS? Qualitative investigations of these sorts of questions are needed.

Despite these qualifications, the concept-focusing strategy used in this study appeared to have a considerable effect on the judgments of the panelists and warrants further study.

## References

Angoff, W. H. (1971). Scales, norms, and equivalent scores. In R. L. Thorndike (Ed.), Educational Measurement, (2nd ed.) (pp. 508-600), Washington, DC: American Council on Education.

Bejar, I. I. (1983). Subject matter experts' assessment of item statistics. Applied Psychological Measurement, 7, 303-310.

Goodwin, L. (In press). Relationships between observed item difficulty levels and Angoff minimum passing levels (MPLs) for a group of "borderline" examinees. Applied Measurement in Education.

Impara, J. C., & Plake, B. S. (1998). Teachers' ability to estimate item difficulty: A test of the assumptions in the Angoff standard setting method. Journal of Educational Measurement, 35, 71-83.

Impara, J. C., & Plake, B. S. (1997). Standard setting: An alternative approach. Journal of Educational Measurement, 34, 355-368.

Kane, M. (1994). Validating the performance standards associated with passing scores. Review of Educational Research, 64, 425-461.

Livingston, S. A. (1995). Standards for reporting the educational achievement of groups. Joint Conference on Standard Setting for Large Scale Assessments, Proceedings Volume II. Washington, DC: National Assessment Governing Board, National Center for Educational Statistics.

Lorge, I., & Kruglov, L. K. (1953). The improvement of estimates of test difficulty. Educational and Psychological Measurement, 13, 34-46.

Psychological Corporation (1993). Metropolitan achievement test: 7th Edition. San Antonio, TX: Author.

Thorndike, R. L. (1982). Item and score conversion by pooled judgment. In Holland, P. W. & Rubin, D. B. (Eds.) Test equating. New York: Academic Press.

Table 1. Actual Means and Standard Deviations for the Borderline and Typical Student and for Rounds 1 and 2 Cut Scores Control and Experimental Group Judges.

| | | Mean | SD |
|---|---|---|---|
| Actual performance | Borderline student | 34.08 | 12.24 |
| | Typical student | 42.55 | 12.79 |

| Cut scores | | Group | | | |
|---|---|---|---|---|---|
| | | Experimental | | Control | |
| Round | | Mean | SD | Mean | SD |
| 1 | Borderline student | 26.35 | 10.10 | 41.17 | 5.57 |
| | Typical student | Unknown[a] | | NA[b] | NA[b] |
| 2 | Borderline student | 28.25 | 6.46 | 36.08 | 6.40 |
| | Typical student | 54.34 | Unknown[a] | NA[b] | NA[b] |
| Absolute Difference | Borderline student | -5.83 | | 2.02 | |
| | Typical student | 11.98 | | NA[b] | NA[b] |
| RMSE Round 2 | Borderline student | .253 (under actual) | | .239 (over actual) | |
| | Typical student | .238 (over actual) | | | |

[a]Server crashed and data were lost. Data reported are all that were recovered or saved on a different drive prior to the crash.
[b] These values not estimated by this group of judges.

Table 2. Correlations between judges item cut scores and actual item difficulties for borderline students and for all students.

| Round | | Group: Experimental | Control |
|---|---|---|---|
| 1 | Borderline students | .715 | .789 |
| | Typical students | Unknown[a] | NA |
| 2 | Borderline students | .867 | .608 |
| | Typical students | .793 | NA[b] |

[a]Server crashed and data were lost. Data reported are all that were recovered or saved on a different drive prior to the crash.
[b] These values not estimated by this group of judges.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

ERIC®

TM028378

# REPRODUCTION RELEASE

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| Title: Utility of a Concept-Focusing Strategy on Judgmental Standard Setting Results | |
| Author(s): James C. Impara, Barbara S. Plake, Melody Hertzog, Gerald Giraud, Robert Spies | |
| Corporate Source: | Publication Date: April, 1998 |

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign at the bottom of the page.

| The sample sticker shown below will be affixed to all Level 1 documents | The sample sticker shown below will be affixed to all Level 2A documents | The sample sticker shown below will be affixed to all Level 2B documents |
|---|---|---|
| PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 1 | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2A | PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY ____ Sample ____ TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) 2B |
| Level 1 | Level 2A | Level 2B |
| ☒ | ☐ | ☐ |
| Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) *and* paper copy. | Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only | Check here for Level 2B release, permitting reproduction and dissemination in microfiche only |

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

*I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.*

**Sign here, → please**

| | | |
|---|---|---|
| Signature: James C. Impara | Printed Name/Position/Title: James C. Impara Director, Buros Inst. for Assessment Consult. & Outreach | |
| Organization/Address: Buros Center for Testing 135 Bancroft Hall, Lincoln, NE 68588-0352 | Telephone: 402-472-8804 | FAX: 402-472-6207 |
| | E-Mail Address: jimpara@unl.edu | Date: 4-8-98 |

(over)

ERIC® Clearinghouse on Assessment and Evaluation

University of Maryland
1129 Shriver Laboratory
College Park, MD 20742-5701

*Tel: (800) 464-3742*
*(301) 405-7449*
*FAX: (301) 405-8134*
*ericae@ericae.net*
*http://ericae.net*

March 20, 1998

Dear AERA Presenter,

Congratulations on being a presenter at AERA[1]. The ERIC Clearinghouse on Assessment and Evaluation invites you to contribute to the ERIC database by providing us with a printed copy of your presentation.

Abstracts of papers accepted by ERIC appear in *Resources in Education (RIE)* and are announced to over 5,000 organizations. The inclusion of your work makes it readily available to other researchers, provides a permanent archive, and enhances the quality of *RIE*. Abstracts of your contribution will be accessible through the printed and electronic versions of *RIE*. The paper will be available through the microfiche collections that are housed at libraries around the world and through the ERIC Document Reproduction Service.

We are gathering all the papers from the AERA Conference. We will route your paper to the appropriate clearinghouse. You will be notified if your paper meets ERIC's criteria for inclusion in *RIE*: contribution to education, timeliness, relevance, methodology, effectiveness of presentation, and reproduction quality. You can track our processing of your paper at http://ericae.net.

Please sign the Reproduction Release Form on the back of this letter and include it with **two** copies of your paper. The Release Form gives ERIC permission to make and distribute copies of your paper. It does not preclude you from publishing your work. You can drop off the copies of your paper and Reproduction Release Form at the **ERIC booth (424)** or mail to our attention at the address below. Please feel free to copy the form for future or additional submissions.

Mail to:

AERA 1998/ERIC Acquisitions
University of Maryland
1129 Shriver Laboratory
College Park, MD 20742

This year ERIC/AE is making a **Searchable Conference Program** available on the AERA web page (http://aera.net). Check it out!

Sincerely,

Lawrence M. Rudner

Lawrence M. Rudner, Ph.D.
Director, ERIC/AE

[1]If you are an AERA chair or discussant, please save this form for future use.

CUA

The Catholic University of America